Adopted·
28 July 2015

# OECD GUIDELINE FOR THE TESTING OF CHEMICALS

## The Larval Amphibian Growth and Development Assay (LAGDA)

### INTRODUCTION

1. The need to develop and validate an assay capable of identifying and characterizing the adverse consequences of exposure to toxic chemicals in amphibians, originates from concerns that environmental levels of chemicals may cause adverse effects in both humans and wildlife. The test guideline of the Larval Amphibian Growth and Development Assay (LAGDA) describes a toxicity test with an amphibian species that considers growth and development from fertilization through the early juvenile period. It is an assay (typically 16 weeks) that assesses early development, metamorphosis, survival, growth, and partial reproductive maturation. It also enables measurement of a suite of other endpoints that allows for diagnostic evaluation of suspected endocrine disrupting chemicals (EDCs) or other types of developmental and reproductive toxicants. The method described in this guideline is derived from validation work on African clawed frog (*Xenopus laevis*) by the U.S. Environmental Protection Agency (U.S. EPA) with supporting work in Japan (1). Although other amphibian species may be adapted to a growth and developmental test protocol with ability to determine genetic sex being an important component, the specific methods and observational endpoints detailed in this guideline are applicable to *Xenopus laevis* alone.

2. The LAGDA serves as a higher tier test with an amphibian for collecting more comprehensive concentration-response information on adverse effects suitable for use in hazard identification and characterization, and in ecological risk assessment. The assay fits at Level 4 of the OECD Conceptual Framework on Endocrine Disrupters Testing and Assessment, where *in vivo* assays also provide data on adverse effects on endocrine relevant endpoints (2). The general experimental design entails exposing *X. laevis* embryos at Nieuwkoop and Faber (NF) stage 8-10 (3) to a minimum of four different concentrations of test chemical (generally spaced at not less than half-logarithmic intervals) and control(s) until 10 weeks after the median time to NF stage 62 in the control, with one interim sub-sample at NF stage 62 ($\leq$ 45 post fertilization; usually around 45 days (dpf). There are four replicates in each test concentration with eight replicates for the control. Endpoints evaluated during the course of the exposure (at the interim sub-sample and final sample at completion of the test) include those indicative of generalized toxicity: mortality, abnormal behavior, and growth determinations (length and weight), as well as endpoints designed to characterize specific endocrine toxicity modes of action targeting oestrogen, androgen or thyroid-mediated physiological processes. The method gives primary emphasis to potential population relevant effects (namely, adverse impacts on survival, development, growth and reproductive development) for the

calculation of a No Observed Effect Concentration (NOEC) or an Effect Concentration causing x% change (ECx) in the endpoint measured. Although it should be noted that ECx approaches are rarely suitable for large studies of this type where increasing the number of test concentrations to allow for determination of the desired ECx may be impractical. It should also be noted that the method does not cover the reproductive phase itself. Definitions used in this Test Guideline are given in Annex 1.

**INITIAL CONSIDERATIONS AND LIMITATIONS**

3. Due to the limited number of chemicals tested and laboratories involved in the validation of this rather complex assay, especially inter-laboratory reproducibility is not documented with experimental data so far, it is anticipated that when a sufficient number of studies is available to ascertain the impact of this new study design, the Test guideline will be reviewed and if necessary revised in light of experience gained. The LAGDA is an important assay to address potential contributors to amphibian population declines by evaluating the effects from exposure to chemicals during the sensitive larval stage, where effects on survival and development, including normal development of reproductive organs, may adversely affect populations.

4. The test is designed to detect an apical effect(s) resulting from both endocrine and non-endocrine mechanisms, and includes diagnostic endpoints which are partly specific to key endocrine modalities. It should be noted that until the LAGDA was developed, no validated assay existed that served this function for amphibians.

5. Before beginning the assay, it is important to have information about the physicochemical properties of the test chemical, particularly to allow the production of stable chemical solutions. It is also necessary to have an adequately sensitive analytical method for verifying test chemical concentrations. Over a duration of approximate 16 weeks, the assay requires a total number of 480 animals, i.e., *X. laevis* embryos, (or 640 embryos, if a solvent control is used) to ensure sufficient power of the test for the evaluation of population-relevant endpoints such as growth, development and reproductive maturation.

6. Before use of the Test Guideline for regulatory testing of a mixture, it should be considered whether it will provide acceptable results for the intended regulatory purpose. Furthermore, this assay does not evaluate fecundity directly, so it may not be applicable for use at a more advanced stage than Level 4 of the OECD Conceptual Framework.

**SCIENTIFIC BASIS FOR THE TEST METHOD**

7. Much of our current understanding of amphibian biology has been obtained using the laboratory model species *X. laevis*. This species can be routinely cultured in the laboratory, ovulation can be induced using human chorionic gonadotropin (hCG) and animal stocks are readily available from commercial breeders.

8 Like all vertebrates, reproduction in amphibians is under the control of the hypothalamic pituitary gonadal (HPG) axis (4). Oestrogens and androgens are mediators of this endocrine system, directing the development and physiology of sexually-dimorphic tissues. There are three distinct phases in the life cycle of amphibians when this axis is especially active: (1) gonadal differentiation during larval development, (2) development of secondary sex characteristics and gonadal maturation during the juvenile phase and (3) functional reproduction of adults. Each of these three developmental windows are likely susceptible to endocrine perturbation by certain chemicals such as estrogens and androgens, ultimately leading to a loss

of reproductive fitness by the organisms.

9. The gonads begin development at NF stage 43, when the bipotential genital ridge first develops. Differentiation of the gonads begins at NF stage 52 when primordial germ cells either migrate to medullary tissue (males) or remain in the cortical region (females) of the developing gonads (3). This process of sexual differentiation of the gonads was first reported to be susceptible to chemical alteration in *Xenopus* in the 1950's (5) (6). Exposure of tadpoles to estradiol during this period of gonad differentiation results in sex reversal of males that when raised to adulthood are fully functional females (7) (8). Functional sex reversal of females into males is also possible and has been reported following implantation of testis tissue in tadpoles (9). However, although exposure to an aromatase inhibitor also causes functional sex reversal in *X. tropicalis* (10), this has not been shown to occur in *X. laevis*. Historically, toxicant effects on gonadal differentiation have been assessed by histological examination of the gonads at metamorphosis and sex reversal could only be determined by analysis of sex ratios. Until recently, there had been no means to directly determine the genetic sex of *Xenopus*. However, recent establishment of sex linked markers in *X. laevis* make it possible to determine genetic sex and allows for the direct identification of sex reversed animals (11).

10. In males, juvenile development proceeds as blood levels of testosterone increase corresponding with the development of secondary sex characteristics as well as testis development. In females, estradiol is produced by the ovaries resulting in the appearance of vitellogenin (VTG) in the plasma, vitellogenic oocytes in the ovary and the development of oviducts (12). Oviducts are female secondary sex characteristics that function in oocyte maturation during reproduction. Jelly coats are applied to the outside of oocytes as they pass through the oviduct and collect in the ovisac, ready for fertilization. Oviduct development appears to be regulated by oestrogens as development correlates with blood estradiol levels in *X. laevis* (13) and *X. tropicalis* (12). The development of oviducts in males following exposure to polychlorinated biphenyl compounds (14) and 4-*tert*-octylphenol (15) has been reported.

## PRINCIPLE OF THE TEST

11. The test design entails exposing *X. laevis* embryos at NF stage 8-10 via the water route to four different concentrations of test chemical as well as control(s) until 10 weeks after the median time to NF stage 62 in the control with one interim sub-sample at NF stage 62. While it may also be possible to dose highly hydrophobic chemicals via the feed, there has been little experience using this exposure route in this assay to date. There are four replicates in each test concentration with eight replicates for each control used. Endpoints evaluated during the course of the exposure include those indicative of generalized toxicity (i.e., mortality, abnormal behavior and growth determinations (length and weight)), as well as endpoints designed to characterize specific endocrine toxicity modes of action targeting oestrogen-, androgen-, or thyroid-mediated physiological processes (i.e. thyroid histopathology, gonad and gonad duct histopathology, abnormal development, plasma vitellogenin (optional), and genotypic/phenotypic sex ratios) .

## TEST VALIDITY CRITERIA

12. The following criteria for test validity apply:

- The dissolved oxygen concentration should be ≥ 40% of air saturation value throughout the test;
- The water temperature should be in the range of 21 ± 1 °C and the inter-replicate and the inter-treatment differentials should not exceed 1.0 °C;
- pH of the test solution should be maintained between 6.5 and 8.5, and the inter-replicate and the inter-treatment differentials should not exceed 0.5;

- Evidence should be available to demonstrate that the concentrations of the test chemical in solution have been satisfactorily maintained within ± 20% of the mean measured values;
- Mortality over the exposure period should be ≤ 20% in each replicate in the controls;
- ≥ 70% viability in the spawn chosen to start the study;
- The median time to NF stage 62 of the controls should be ≤ 45 days.
- The mean weight of test organisms at NF stage 62 and at the termination of the assay in controls and solvent controls (if used) should reach 1.0 ± 0.2 and 11.5 ± 3 g, respectively.

13. While not a validity criterion, it is recommended that at least three treatment levels with three uncompromised replicates be available for analysis. Excessive mortality, which compromises a treatment, is defined as > 4 mortalities (> 20%) in 2 or more replicates that cannot be explained by technical error. At least three treatment levels without obvious overt toxicity should be available for analysis. Signs of overt toxicity may include, but are not limited to, floating on the surface, lying on the bottom of the tank, inverted or irregular swimming, lack of surfacing activity, and being nonresponsive to stimuli, morphological abnormalities (*e.g.*, limb deformities), hemorrhagic lesions, and abdominal oedema.

14. In case a deviation from the test validity criteria is observed, the consequences should be considered in relation to the reliability of the test results, and these deviations and considerations should be included in the test report.

**DESCRIPTION OF THE METHODS**

***Apparatus***

15. Normal laboratory equipment and especially the following:

(a) temperature controlling apparatus (e.g., heaters or coolers adjustable to 21 ± 1 °C);
(b) thermometer;
(c) binocular dissection microscope and dissection tools;
(d) digital camera with at least 4 megapixel resolution and micro function (if needed);
(e) analytical balance capable of measuring to 0.001 mg or 1 μg;
(f) dissolved oxygen meter and pH meter;
(g) light intensity meter capable of measuring in lux units.

***Water***

*Source and quality*

16. Any dilution water that is locally available (e.g. spring water or charcoal-filtered tap water) and permits normal growth and development of *X. laevis* can be used, and evidence of normal growth in this water should be available. Because local water quality can differ substantially from one area to another, analysis of water quality should be undertaken, particularly if historical data on the utility of the water for raising amphibian larvae is not available. Measurements of heavy metals (e.g. Cu, Pb, Zn, Hg, Cd, Ni), major anions and cations (e.g. Ca, Mg, Na, K, Cl, $SO_4$), pesticides, total organic carbon and suspended solids should be made before testing begins and/or, for example, every six months where a dilution water is known to be relatively constant in quality. Some chemical characteristics of acceptable dilution water are listed in ANNEX 2.

*Iodide concentration in test water*

17. In order for the thyroid gland to synthesize thyroid hormones to support normal metamorphosis, sufficient iodide should to be available to the larvae through a combination of aqueous and dietary sources. Currently, there are no empirically derived guidelines for minimum iodide concentrations in either food or water to ensure proper development. However, iodide availability may affect the responsiveness of the thyroid system to thyroid active agents and is known to modulate the basal activity of the thyroid gland which deserves attention when interpreting the results from thyroid histopathology. Based on previous work, successful performance of the assay has been demonstrated when dilution water iodide ($I^-$) concentrations range between 0.5 and 10 µg/L. Ideally, the minimum iodide concentration in the dilution water throughout the test should be 0.5 µg/L (added as the sodium or potassium salt). If the test water is reconstituted from deionized water, iodine should be added at a minimum concentration of 0.5 µg/L. The measured iodide concentrations from the test water (i.e., dilution water) and the supplementation of the test water with iodine or other salts (if used) should be reported. Iodine content may also be measured in food(s) in addition to test water.

***Exposure system***

18. The test was developed using a flow-through diluter system. The system components should have water-contact components of glass, stainless steel, and/or other chemically inert materials. Exposure tanks should be glass or stainless steel aquaria and tank usable volume should be between 4.0 and 10.0 L (minimum water depth of 10 to 15 cm). The system should be capable of supporting all exposure concentrations, a control, and a solvent control, if necessary, with four replicates per treatment and eight in the controls. The flow rate to each tank should be constant in consideration of both the maintenance of biological conditions and chemical exposure. It is recommended that flow rates should be appropriate (***e.g.***, at least 5 tank turnovers per day) to avoid chemical concentration declines due to metabolism by both the test organisms and aquatic microorganisms present in the aquaria or abiotic routes of degradation (hydrolysis, photolysis) or dissipation (volatilization, sorption). The treatment tanks should be randomly assigned to a position in the exposure system to reduce potential positional effects, including slight variations in temperature, light intensity, *etc*. Further information on setting up flow-through exposure systems can be obtained from the ASTM Standard Guide for Conducting Acute Toxicity Tests on Test Materials with Fishes, Macroinvertebrates, and Amphibians (16).

***Chemical delivery: preparation of test solutions***

19. To make test solutions in the exposure system, stock solution of the test chemical should be dosed into the exposure system by an appropriate pump or other apparatus. The flow rate of the stock solution should be calibrated in accordance with analytical confirmation of the test solutions before the initiation of exposure, and checked volumetrically periodically during the test. The test solution in each chamber should be renewed at a minimum of 5 volume renewals/day.

20. The method used to introduce the test chemical to the system can vary depending on its physicochemical properties. Therefore, prior to the test, baseline information about the chemical that is relevant to determining its testability should be obtained. Useful information about test chemical-specific properties include the structural formula, molecular weight, purity, stability in water and light, pKa and Kow, water solubility (preferably in the test medium) and vapour pressure as well as results of a test for ready biodegradability (OECD TG 301 (17) or TG 310 (18)). Solubility and vapour pressure can be used to calculate Henry's law constant, which will indicate whether losses due to evaporation of the test chemical may occur. Conduct of this test without the information listed above should be carefully considered as the study design will be dependent on the physicochemical properties of the test chemical and, without these data test results may be difficult to interpret or meaningless. A reliable analytical method for the quantification of the test chemical in the test solutions with known and reported accuracy and limit of detection should be available. Water soluble test chemicals can be dissolved in aliquots of dilution water at

a concentration which allows delivery at the target test concentration in a flow-through system. Chemicals which are liquid or solid at room temperature and moderately soluble in water may require liquid:liquid or liquid:solid (*e.g.*, glass wool column) saturators (19). While it may also be possible to dose very hydrophobic test chemicals via the feed, there has been little experience using that exposure route in this assay.

21. Test solutions of the chosen concentrations are prepared by dilution of a stock solution. The stock solution should preferably be prepared by simply mixing or agitating the test chemical in dilution water by mechanical means (e.g. stirring and/or ultrasonication). Saturation columns/systems or passive dosing methods (20) can be used for achieving a suitably concentrated stock solution. The preference is to use a co-solvent-free test system; however, different test chemicals will possess varied physicochemical properties that will likely require different approaches for preparation of chemical exposure water. All efforts should be made to avoid solvents or carriers because: (1) certain solvents themselves may result in toxicity and/or undesirable or unexpected responses, (2) testing chemicals above their water solubility (as can frequently occur through the use of solvents) can result in inaccurate determinations of effective concentrations, (3) the use of solvents in longer-term tests can result in a significant degree of "biofilming" associated with microbial activity which may impact environmental conditions as well as the ability to maintain exposure concentrations and (4) the absence of historical data that demonstrate that the solvent does not influence the outcome of the study, use of solvents requires a solvent control treatment which has significant animal welfare implications as additional animals are required to conduct the test. For difficult to test chemicals, a solvent may be employed as a last resort, and the OECD Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures should be consulted (21) to determine the best method. The choice of solvent will be determined by the chemical properties of the test chemical and the availability of historical control data on the solvent. In the absence of historical data, the suitability of a solvent should be determined prior to conducting the definitive study. In the event that use of a solvent is unavoidable, and microbial activity (biofilming) occurs, recommend recording/reporting of the biofilming per tank (at least weekly) throughout the test. Ideally, the solvent concentration should be kept constant in the solvent control and all test treatments. If the concentration of solvent is not kept constant, the highest concentration of solvent in the test treatment should be used in the solvent control. In cases where a solvent carrier is used, maximum solvent concentrations should not exceed 100 µl/L or 100 mg/L (21), and it is recommended to keep solvent concentration as low as possible (*e.g*, $\leq$20 µl/L) to avoid potential effects of the solvent on endpoints measured (22).

***Test animals***

*Test species*

22. The test species is *X. laevis* because this is: (1) routinely cultured in laboratories worldwide, (2) easily obtainable through commercial suppliers and (3) capable of having its genetic sex determined.

*Adult care and breeding*

23. Appropriate care and breeding of *X. laevis* is described by a standardized guideline (23). Housing and care of *X. laevis* are also described by Read (24). To induce breeding, three to five pairs of adult females and males are injected intraperitoneally with human chorionic gonadotropin (hCG). Female and male specimens are injected with e.g., approximately 800-1000 IU and 500-800 IU, respectively, of hCG dissolved in 0.6-0.9% saline solution (or frog Ringer's solution, an isotonic saline for use with amphibians; www.hermes.mbl.edu/biologicalbulletin/compendium/comp-RGR.html ). Injection volumes should be about 10 µl/g body weight (~1000 µl). Afterwards, induced breeding pairs are held in large tanks, undisturbed and under static conditions to promote amplexus. The bottom of each breeding tank should have a false bottom of stainless steel mesh (e.g., 1.25 cm openings) which permits the eggs to fall to

the bottom of the tank. Frogs injected with hCG in the late afternoon will usually deposit most of their eggs by mid-morning of the next day. After a sufficient quantity of eggs is released and fertilized, adults should be removed from the breeding tanks. Eggs are then collected and jelly coats are removed by L-cysteine treatment (23). A 2% L-cysteine solution should be prepared and pH adjusted to 8.1 with 1 M NaOH. This 21 °C solution is added to a 500 mL Erlenmeyer flask containing the eggs from a single spawn and swirled gently for one to two minutes and then rinsed thoroughly 6-8 times with 21 °C culture water. The eggs are then transferred to a crystallizing dish and determined to be > 70% viable with minimal abnormalities in embryos exhibiting cell division.

**TEST DESIGN**

***Test concentrations***

24. It is recommended to use a minimum of four chemical concentrations and appropriate controls (including solvent controls, if necessary). Generally, a concentration separation (spacing factor) not exceeding 3.2 is recommended.

25. For the purposes of this test, results from existing amphibian studies should be used to the extent possible in determining the highest test concentration so as to avoid concentrations that are overtly toxic. Information from, for example, quantitative structure-activity relationships, read across and data from existing amphibian studies such as the Amphibian Metamorphosis Assay, TG231 (25) and the Frog Embryo Teratogenesis Assay - *Xenopus* (23) and/or fish tests such as OECD TG229, TG234 and TG236 (26) (27) (28) may contribute toward setting this concentration. Prior to running the LAGDA a range finding experiment may be conducted. It is recommended that the range-finding exposure is initiated within 24 hours of fertilization and continued for 7-14 days (or more, if needed), and the test concentrations are set such that the intervals between test concentrations are no greater than a factor of 10. The results of the range finding experiment should serve to set the highest test concentration in the LAGDA. Note that if a solvent has to be used, then the suitability of the solvent (i.e. whether it may have an impact on the outcome of the study) could be determined as part of the range finding study.

***Replicates within treatment groups and controls***

26. A minimum of four replicate tanks per test concentration and a minimum of eight replicates for the controls (and solvent control, if needed) should be used (i.e., the number of replicates in the control and any solvent control should be twice as large as the number of replicates of each treatment group, to ensure appropriate statistical power). Each replicate should contain no more than 20 animals. While using the test method in this guideline, the minimum number of animals processed would be 15 (5 for NF stage 62 sub-sample and 10 juveniles). However, additional animals are added to each replicate to factor in the possibility for mortality while maintaining the critical number of 15.

**PROCEDURE**

***Assay overview***

27. The assay is initiated with newly spawned embryos (NF stage 8-10) and continues into juvenile development. Animals are examined daily for mortality and any sign of abnormal behavior. At NF stage 62, a larval sub-sample (up to 5 animals per replicate) is collected and various endpoints are examined (**Table 1**). After all animals have reached NF stage 66, i.e. completion of metamorphosis (or after 70 days from the assay initiation, whichever comes first), a cull is carried out at random (but without sub-sampling) to reduce the number of animals (10 per tank) (see paragraph 43), and the remaining animals continue exposure until 10 weeks after the median time to NF stage 62 in the control. At test termination (juvenile sampling) additional measurements are made (**Table 1**).

***Exposure conditions***

28. A complete summary of test parameters can be found in ANNEX 3. During the exposure period, dissolved oxygen, temperature, and pH of test solutions should be measured daily. Conductivity, alkalinity, and hardness are measured once a month. For the water temperature of test solutions, the inter-replicate and inter-treatment differentials (within one day) should not exceed 1.0 °C. Also, for pH of test solutions, the inter-replicate and inter-treatment differentials should not exceed 0.5.

29. The exposure tanks may be siphoned on a daily basis to remove uneaten food and waste products, being careful to avoid cross-contamination of tanks. Care should be used to minimize stress and trauma to the animals, especially during movement, cleaning of aquaria, and manipulation. Stressful conditions/activities should be avoided such as loud and/or incessant noise, tapping on aquaria, vibrations in the tank.

***Duration of exposure to the test chemical***

30. The exposure is initiated with newly spawned embryos (NF stage 8-10) and continued until ten weeks after the median time to NF stage 62 (≤ 45 days from the assay initiation) in control group. Generally, the duration of the LAGDA is 16 weeks (maximum 17 weeks).

***Initiation of assay***

31. Parent animals used for the initiation of the assay should have previously been shown to produce offspring that can be genetically sexed (ANNEX 5). After spawning of adults, embryos are collected, cysteine-treated to remove the jelly coat and screened for viability (ASTM, 2004). Cysteine treatment allows the embryos to be handled during screening without sticking to surfaces. Screening takes place under a dissecting microscope using an appropriately sized eye dropper to remove non-viable embryos. It is preferred that a single spawn resulting in greater than 70% viability be used for the test. Embryos at NF stage 8-10 are randomly distributed into exposure treatment tanks containing an appropriate volume of dilution water until each tank contains 20 embryos. Embryos should be carefully handled during this transfer in order to minimize handling stress and to avoid any injury. At 96 hours post fertilization, the tadpoles should have moved up the water column and begun clinging to the sides of the tank.

***Feeding regime***

32. Feed and feeding rate change during different life stages of *X. laevis* are a very important aspect of the LAGDA protocol. Excessive feeding during the larval phase typically results in increased incidences and severity of scoliosis (ANNEX 8) and should be avoided. Conversely, inadequate feeding during the larval phase results in highly variable developmental rates among controls potentially compromising statistical power or confounding test results. ANNEX 4 provides recommended larval and juvenile diet and feeding regimes for *X. laevis* in flow-through conditions, but alternatives are permissible providing the test organisms grow and develop satisfactorily. It is important to note that if endocrine-specific endpoints are being measured, feed should be free of endocrine-active substances such as soy meal.

*Larval feeding*

33. The recommended larval diet consists of trout starter feeds, *Spirulina* algae discs and goldfish crisps (*e.g.*, TetraFin® flakes, Tetra, Germany) blended together in culture (or dilution) water. This mixture is administered three times daily on weekdays and once daily on weekends. Tadpoles are also fed live brine shrimp, *Artemia* spp., 24-hour-old nauplii, twice daily on weekdays and once daily on the weekends starting on day 8 post-fertilization. The larval feeding, which should be consistent in each test vessel,

should allow appropriate growth and development for test animals in order to ensure reproducibility and transferability of the assay results: (1) the median time to NF stage 62 in controls should be ≤ 45 days and (2) a mean weight within 1.0 ± 0.2 g at NF stage 62 in controls is recommended.

*Juvenile feeding*

34. Once metamorphosis is complete, the feeding regime consists of premium sinking frog food, *e.g.*, Sinking Frog Food -3/32 (Xenopus Express, FL, USA) (ANNEX 4). For froglets (early juveniles), the pellets are briefly run in a coffee grinder, blender or crushed with a mortar and pestle in order to reduce their size. Once juveniles are large enough to consume full pellets, grinding or crushing is no longer necessary. The animals should be fed once per day. The juvenile feeding should allow appropriate growth and development of the organisms: a mean weight within 11.5 ± 3 g in control juveniles at the termination of the assay is recommended.

***Analytical chemistry***

35. Prior to initiation of the assay, the stability of the test chemical (*e.g.*, solubility, degradability, and volatility) and all analytical methods needed should be established *e.g.*, using existing information or knowledge. When dosing via the dilution water, it is recommended that test solutions from each replicate tank concentration be analyzed prior to test initiation to verify system performance. During the exposure period, the concentrations of the test chemical are determined at appropriate intervals, preferably every week for at least one replicate in each treatment group, rotating between replicates of the same treatment group every week. It is recommended that results be based on measured concentrations. However, if concentration of the test chemical in solution has been satisfactorily maintained within ± 20% of the nominal concentration throughout the test, then the results can either be based on nominal or measured values. Also, the coefficient of variation (CV) of the measured test concentrations over the entire test period within a treatment should be maintained at 20% or less in each concentration. When the measured concentrations do not remain within 80-120% of the nominal concentration (for example, when testing highly biodegradable or adsorptive chemicals), the effect concentrations should be determined and expressed relative to the arithmetic mean concentration for flow-through tests.

36. The flow rates of dilution water and stock solution should be checked at appropriate intervals (*e.g.* three times a week) throughout the exposure duration. In the case of chemicals which cannot be detected at some or all of the nominal concentrations, (*e.g.*, due to rapid degradation or adsorption in the test vessels, or by marked chemical accumulation in the bodies of exposed animals), it is recommended that the renewal rate of the test solution in each chamber be adapted to maintain test concentrations as constant as possible.

***Observations and endpoint measurements***

37. The endpoints evaluated during the course of the exposure are those indicative of toxicity including mortality, abnormal behavior such as clinical signs of disease and/or general toxicities, and growth determinations (length and weight), as well as pathology endpoints which may respond to both general toxicity and endocrine modes of action targeting oestrogen-, androgen-, or thyroid-mediated pathways. In addition, plasma VTG concentration may be optionally measured at the termination of the assay. Measurement of VTG may be useful in understanding study results in the context of endocrine mechanisms for suspected EDCs. The endpoints and timing of measurements are summarized in **Table 1**.

**Table 1.** Endpoint overview of the LAGDA.

| Endpoints* | Daily | Interim Sampling (Larval sampling) | Test Termination (Juvenile sampling) |
|---|---|---|---|
| Mortality and abnormalities | X | | |
| Time to NF stage 62 | | X | |
| Histo(patho)logy (thyroid gland) | | X | |
| Morphometrics (growth in weight and length) | | X | X |
| Liver-somatic index (LSI) | | | X |
| Genetic/phenotypic sex ratios | | | X |
| Histopathology (gonads, reproductive ducts, kidney and liver) | | | X |
| Vitellogenin (VTG) (optional) | | | X |

*** All endpoints are analyzed statistically.**

***Mortality and daily observations***

38. All test tanks should be checked daily for dead animals and mortalities recorded for each tank. Dead animals should be removed from the test tank as soon as observed. The developmental stage of dead animals should be categorized as either pre-NF stage 58 (pre-forelimb emergence), NF stage 58-NF stage 62, NF stage 63-NF stage 66 (between NF stage 62 and complete tail absorption), or post-NF stage 66 (post-larval). Mortality rates exceeding 20% may indicate inappropriate test conditions or overtly toxic effects of the test chemical. The animals tend to be most sensitive to non-chemical induced mortality events during the first few days of development after the spawning event and during metamorphic climax. Such mortality could be apparent from the control data.

39. In addition, any observation of abnormal behavior, grossly visible malformations (*e.g.*, scoliosis), or lesions should be recorded. Observations of scoliosis should be counted (incidence) and graded with respect to severity (*e.g.*, not remarkable – NR, minimal – 1, moderate – 2, severe – 3; ANNEX 8). Efforts should be made to ensure that the prevalence of moderate and severe scoliosis is limited (*e.g.*, below 10% in controls) throughout the study, although greater prevalence of control abnormalities would not necessarily be a reason for stopping the test. Normal behavior for larval animals is characterized by suspension in the water column with tail elevated above the head, regular rhythmic tail fin beating, periodic surfacing, operculating, and being responsive to stimuli. Abnormal behaviors would include, for example, floating on the surface, lying on the bottom of the tank, inverted or irregular swimming, lack of surfacing activity, and being nonresponsive to stimuli. For post-metamorphic animals, in addition to the above abnormal behaviors, gross differences in food consumption between treatments should be recorded. Gross malformations and lesions could include morphological abnormalities (*e.g.*, limb deformities), hemorrhagic lesions, abdominal edema, and bacterial or fungal infections, to name a few. The occurrences of lesions on the head of juveniles, just posterior to the nostrils, may be indications of insufficient humidity levels. These determinations are qualitative and should be considered akin to clinical signs of disease/stress and made in comparison to control animals. If the rate of occurrence is greater in exposed tanks than in the controls, then these should be considered as evidence for overt toxicity.

***Larval sub-sampling***

*Outline of larval sub-sampling:*

40. The tadpoles that have reached NF stage 62 should be removed from the tanks and either sampled or moved to the next part of the exposure in a new tank, or physically separated from the remaining tadpoles in the same tank with a divider. Tadpoles are checked daily, and the study day on which an individual tadpole reaches NF stage 62 is recorded. The defining characteristic for use in this assessment is the shape of the head. Once the head has become reduced in size such that it is visually approximately the same width as the trunk of the tadpole and forelimb at the level of the middle of the heart, then that individual would be counted as having attained NF stage 62.

41. The goal is to sample a total of five NF stage 62 tadpoles per replicate tank. This should be performed entirely at random, but decided *a priori*. A hypothetical example of a replicate tank is provided in **Figure 1**. Should there be 20 surviving tadpoles in a particular tank when the first individual reaches NF stage 62, five random numbers should be chosen from 1-20. Tadpole #1 is the first individual to reach NF stage 62 and tadpole #20 is the last individual in a tank to reach NF stage 62. Likewise, if there are 18 surviving larvae in a tank, five random numbers should be chosen from 1-18. This should be performed for every replicate tank when the first individual on test reaches NF stage 62. If there are mortalities during the NF stage 62 sampling, the remaining samples need to be re-randomized based on how many larvae are left <NF stage 62 and how many more samples are needed to reach a total of five samples from that replicate. On the day a tadpole reaches NF stage 62, reference to the prepared sampling chart is made to determine whether that individual is sampled or physically separated from the remaining tadpoles for continued exposure. In the example provided (**Figure 1**), the first individual to reach NF stage 62 (i.e. box #1) is physically separated from the other larvae, continues exposure and the study day on which that individual reached NF stage 62 is recorded. Subsequently, individuals #2 and #3 are treated the same way as #1 and then individual #4 is sampled for growth and thyroid histology (according to this example). This procedure continues until the 20$^{th}$ individual either joins the rest of the post-NF stage 62 individuals or is sampled. The random procedure used must give each organism on test equal probability of being selected. This can be achieved by using any randomizing method, but also requires that each tadpole be netted at some point throughout the NF stage 62 sub-sampling period.

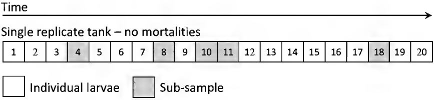


**Figure 1**. Hypothetical example of NF stage 62 sampling regime for a single replicate tank.

42. For the larval sub-sampling, the endpoints obtained are: (1) time to NF stage 62 (i.e., number of days between fertilization and NF stage 62), (2) external abnormalities, (3) morphometrics (*e.g*., weight and length) and (4) thyroid histology.

*Humane killing of tadpoles*

43. The sub-sample of NF stage 62 tadpoles (5 individuals per replicate) should be euthanized by immersion for 30 minutes in appropriate amounts (*e.g.*, 500 mL) of anesthetic solution (*e.g.*, 0.3% solution of MS-222, tricaine methane sulfonate, CAS 886-86-2). MS-222 solution should be buffered with sodium bicarbonate to a pH of approximately 7.0 because unbuffered MS-222 solution is acidic and irritating to frog skin resulting in poor absorption and unnecessary additional stress to the organisms.

44. Using a mesh dip net, a tadpole is removed from the experimental chamber and transported (placed) into the euthanasia solution. The animal is properly euthanized and is ready for necropsy when it is unresponsive to external stimuli such as pinching the hind limb with a pair of forceps.

*Morphometrics (weight and length)*

45. Measurements of wet weight (nearest mg) and snout-to-vent length (SVL) (nearest 0.1 mm) for each tadpole should be made immediately after it becomes non-responsive by anesthesia (**Figure 2a**). Image analysis software may be used to measure SVL from a photograph. Tadpoles should be blotted dry before weighing to remove excess adherent water. After measurements of body size (weight and SVL) are made, any gross morphological abnormalities and/or clinical signs of toxicity such as scoliosis (see ANNEX 8), petechiae and hemorrhage should be recorded or noted, and digital documentation is recommended. Note that petechiae are small red or purple hemorrhages in skin capillaries.

*Tissue Collection and Fixation*

46. For the larval sub-sample, thyroid glands are assessed for histology. The lower torso posterior to the forelimbs is removed and discarded. The trimmed carcass is fixed in Davidson's fixative. The volume of fixative in the container should be at least 10 times the approximate volume of the tissues. Appropriate agitation or circulation of the fixative should be achieved to adequately fix the tissues of interest. All tissues remain in Davidson's fixative for at least 48 hours, but no longer than 96 hours, at which time they are rinsed in deionized water and stored in 10% neutral buffered formalin (1) (29).

*Thyroid histology*

47. Each larval sub-sample (tissues fixed) is histologically assessed for thyroid glands, i.e., diagnosis and severity grading (29) (30).

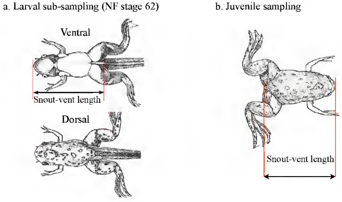


**Figure 2.** Landmarks for measuring snout-vent length for the LAGDA in NF Stage 62 (a) and juvenile frogs (b). The defining characteristics of NF stage 62 (a): the head is the same width as the trunk, the olfactory nerve length is shorter than the diameter of the olfactory bulb (dorsal view), and the forelimbs are at the level of the heart (ventral view). Images adapted from Nieuwkoop and Faber (1994).

***End of larval exposure***

48. Given the initial number of tadpoles, it is expected that there will likely be a small percentage of individuals that do not develop normally and do not complete metamorphosis (NF stage 66) in a reasonable amount of time. The larval portion of the exposure should not exceed 70 days. Any tadpoles remaining at the end of this period should be euthanized (see para 43), their wet weight and SVL measured, staged according to Nieuwkoop and Faber, 1994, and any developmental abnormalities noted

***Cull after NF stage 66***

49. Ten individuals per tank should continue from NF stage 66 (complete tail resorption) until termination of the exposure. Therefore, after all animals have reached NF stage 66 or after 70 days (whichever occurs first), a cull should be conducted. Post NF stage 66 animals that will not continue the exposure should be selected at random.

50. Animals that are not selected for continued exposure are euthanized (see para. 43). Measurements of developmental stage, wet weight and SVL (**Figure 2**b) and a gross necropsy are conducted for each animal. The phenotypic sex (based on gonad morphology) is noted as female, male, or indeterminate.

***Juvenile Sampling***

*Outline of juvenile sampling*

51. The remaining animals continue exposure until 10 weeks after the median time to NF stage 62 in the dilution water (and/or solvent control if relevant) control. At the end of the exposure period, the remaining animals (maximum 10 frogs per replicate) are euthanized, and the various endpoints are measured or evaluated and recorded: (1) morphometrics (weight and length), (2) phenotypic/genotypic sex ratios, (3) liver weight (Liver-Somatic Index), (4) histopathology (gonads, reproductive ducts, liver and kidney) and optionally (5) plasma VTG.

*Humane killing of frogs*

52. The juvenile samples, post-metamorphic frogs, are euthanized by an intraperitoneal injection of anesthetic, *e.g.*, 10% MS-222 in an appropriate phosphate buffered solution. Frogs may be sampled after becoming unresponsive (usually around 2 min after injection, if 10% MS-222 is used in a dosage of 0.01 mL per g of frog). While the juvenile frogs could be immersed in a higher concentration of anesthetic (MS-222), experience has shown that it takes longer for them to be anesthetized using this method and the duration may not be adequate to allow for sampling. Injection provides efficient, fast euthanasia prior to sampling. Sampling should not be started until lack of responsiveness of the frogs has been confirmed to ensure that the animals are dead. If frogs are showing signs of considerable suffering (very severe and death can be reliably predicted) and considered moribund, animals should be anaesthetized and euthanized and treated as mortality for data analysis. When a frog is euthanized due to morbidity, this should be noted and reported. Depending on when the frog is euthanized during the study, retaining the frog for histopathology analysis may be conducted (fixing the frog for possible histopathology).

*Morphometrics (weight and length)*

53. Measurements of wet weight and SVL **(Figure 2b)** are identical to those outlined for the larval sub-sampling.

*Plasma VTG (option)*

54. VTG is a widely accepted biomarker resulting from exposure to oestrogenic chemicals. For the LAGDA, plasma VTG optionally may be measured within juvenile samples (this may be particularly relevant if the test chemical is suspected of being an oestrogen).

55. The euthanized juvenile hind limbs are cut and blood is collected with a heparinized capillary (although alternative blood collection methods, such as cardiac puncture, may be suitable). The blood is expelled into a microcentrifuge tube (*e.g.*, 1.5 mL volume) and centrifuged to obtain plasma. The plasma samples should be stored at -70 °C or below until VTG determination. Plasma VTG concentration can be measured by an enzyme-linked immunosorbent assay (ELISA) method (ANNEX 6), or by an alternative method such as mass spectrometry (31). Species specific antibodies are preferred due to greater sensitivity.

*Genetic sex determination*

56. The genetic sex of each juvenile frog is assessed based on the markers developed by Yoshimoto *et al.* (11). To determine the genetic sex, a portion (or whole) of one hind limb (or any other tissue) removed during dissection is collected and stored in a microcentrifuge tube (tissue samples from frogs can be obtained from any tissue). Tissue can be stored at -20°C or below until isolation of deoxyribose nucleic acid (DNA). The isolation of DNA from tissues can be performed with commercially available kits and analysis for presence or absence of the marker is done by a polymerase chain reaction (PCR) method (ANNEX 5). Generally, the concordance between histological sex and genotype across control animals at the juvenile sampling time point in control groups is more than 95%.

*Tissue collection and fixation for histopathology*

57. Gonads, reproductive ducts, kidneys and livers are collected for histological analysis during the final sampling. The abdominal cavity is opened, and the liver is dissected out and weighed. Next, the digestive organs (*e.g.*, stomach, intestines) are carefully removed from the lower abdomen to reveal the gonads, kidneys and reproductive ducts. Any gross morphological abnormalities in the gonads should be noted. Finally, the hind limbs should be removed if they have not previously been removed for blood collection. Collected livers and the carcass with the gonads left *in situ* should be immediately placed into Davidson's fixative. The volume of fixative in the container should be at least 10 times the approximate volume of the tissues. All tissues remain in Davidson's fixative for at least 48 hours, but no longer than 96 hours at which time they are rinsed in de-ionized water and stored in 10% neutral buffered formalin (1) (29).

*Histopathology*

58. Each juvenile sample is evaluated histologically for pathology in the gonads, reproductive ducts, kidneys and liver tissue, i.e., diagnosis and severity grading (32). The gonad phenotype is also derived from this evaluation (*e.g.*, ovary, testis, intersex), and together with individual genetic sex measurements, these observations can be used to calculate phenotypic/genotypic sex ratios.

## DATA REPORTING

***Statistical analysis***

59. The LAGDA generates three forms of data to be statistically analyzed: (1) quantitative continuous data (weight, SVL, LSI, VTG), (2) time-to-event data for developmental rates (i.e., days to NF stage 62 from assay initiation) and (3) ordinal data in the form of severity scores or developmental stages from histopathology evaluations.

60. It is recommended that the test design and selection of statistical test permit adequate power to detect changes of biological importance in endpoints where a NOEC or ECx is to be reported. Statistical analyses of the data (generally, replicate mean basis) should preferably follow procedures described in the document Current Approaches in the Statistical Analysis of Ecotoxicity Data: A Guidance to Application (33). ANNEX 7 of this guideline provides the recommended statistical analysis decision tree and guidance for the treatment of data and in the choice of the most appropriate statistical test or model to use in the LAGDA.

61. The data from juvenile sampling (*e.g.*, growth, LSI) should be analyzed for each genotypic sex separately since genotypic sex is determined for all frogs.

***Data analysis considerations***

*Use of compromised replicates and treatments*

62. Replicates and treatments may become compromised due to excess mortality from overt toxicity, disease, or technical error. If a treatment is compromised from disease or technical error, there should be three uncompromised treatments with three uncompromised replicates available for analysis. If overt toxicity occurs in the high treatment(s), it is preferable that at least three treatment levels with three uncompromised replicates are available for analysis (consistent with the Maximum Tolerated Concentration approach for OECD test guidelines (34). In addition to mortality, signs of overt toxicity may include behavioural effects (e.g. floating on the surface, lying on the bottom of the tank, inverted or irregular swimming, lack of surfacing activity), morphological lesions (e.g. haemorrhagic lesions, abdominal oedema) or inhibition of normal feeding responses when compared qualitatively to control animals.

*Solvent control*

63. At the termination of the test, an evaluation of the potential effects of the solvent (if used) should be performed. This is done through a statistical comparison of the solvent control group and the dilution water control group. The most relevant endpoints for consideration in this analysis are growth determinants (weight and length), as these can be affected through generalized toxicities. If statistically significant differences are detected in these endpoints between the dilution water control and solvent control groups, best professional judgment should be used to determine if the validity of the test is compromised. If the two controls differ, the treatments exposed to the chemical should be compared to the solvent control unless it is known that comparison to the dilution water control is preferred. If there is no statistically significant difference between the two control groups it is recommended that the treatments exposed to the test chemical are compared with the pooled (solvent and dilution water control groups), unless it is known that comparison to either the dilution-water or solvent control group only is preferred.

***Test report***

64. The test report should include the following:

*Test chemical:*

- Physical nature and, where relevant, physicochemical properties;
- Mono-constituent substance:

  physical appearance, water solubility, and additional relevant physicochemical properties;

  chemical identification, such as IUPAC or CAS name, CAS number, SMILES or InChI code, structural formula, purity, chemical identity of impurities as appropriate and practically feasible, etc. (including the organic carbon content, if appropriate).

- Multi-constituent substance, UVBCs and mixtures:

  characterized as far as possible by chemical identity (see above), quantitative occurrence and relevant physicochemical properties of the constituents.

*Test species:*

- Scientific name, strain if available, source and method of collection of the fertilized eggs and subsequent handling.
- Incidence of scoliosis in historical controls for the stock culture used.

*Test conditions:*

- Photoperiod(s);
- Test design (*e.g.*, chamber size, material and water volume, number of test chambers and replicates, number of test organisms per replicate);
- Method of preparation of stock solutions and frequency of renewal (the solubilizing agent and its concentration should be given, when used);
- Method of dosing the test chemical (*e.g.*, pumps, diluting systems);
- The recovery efficiency of the method and the nominal test concentrations, the limit of quantification, the means of the measured values and their standard deviations in the test vessels

and the method by which these were attained and evidence that the measurements refer to the concentrations of the test chemical in true solution;

- Dilution water characteristics: pH, hardness, temperature, dissolved oxygen concentration, residual chlorine levels (if measured), total iodine, total organic carbon (if measured), suspended solids (if measured), salinity of the test medium (if measured) and any other measurements made;
- The nominal test concentrations, the means of the measured values and their standard deviations;
- Water quality within test vessels, pH, temperature (daily) and dissolved oxygen concentration;
- Detailed information on feeding (*e.g.*, type of foods, source, amount given and frequency).

*Results:*

- Evidence that controls met the validity criteria;
- Data for the control (plus solvent control when used) and the treatment groups as follows: mortality and abnormality observed, time to NF stage 62, thyroid histology assessment (larval sample only), growth (weight and length), LSI (juvenile sample only), genetic/phenotypic sex ratios (juvenile sample only), histopathology assessment results for gonads, reproductive ducts, kidney and liver (juvenile sample only) and plasma VTG (juvenile sample only, if performed);
- Approach for the statistical analysis and treatment of data (statistical test or model used);
- No observed effect concentration (NOEC) for each response assessed;
- Lowest observed effect concentration (LOEC) for each response assessed (at $\alpha = 0.05$); ECx for each response assessed, if applicable, and confidence intervals (*e.g.*, 95%) and a graph of the fitted model used for its calculation, the slope of the concentration-response curve, the formula of the regression model, the estimated model parameters and their standard errors.
- Any deviation from the guideline and deviations from the acceptance criteria, and considerations of potential consequences on the outcome of the test.

65. For the results of endpoint measurements, mean values and their standard deviations (on both replicate and concentration basis, if possible) should be presented.

66. Median time to NF stage 62 in controls should be calculated and presented as the mean of replicate medians and their standard deviation. Likewise, for treatments, a treatment median should be calculated and presented as the mean of replicate medians and their standard deviation.

## LITERATURE

1. U.S. Environmental Protection Agency. (2013). Validation of the Larval Amphibian Growth and Development Assay: Integrated Summary Report.

2. OECD. (2012a) Guidance Document on Standardised Test Guidelines for Evaluating Endocrine Disrupters.Environment, Health and Safety Publications, Series on testing and assessment (No. 150.) Organisation for Economic Cooperation and Development, Paris.

3. Nieuwkoop PD and Faber J. (1994). Normal Table of *Xenopus laevis* (Daudin). Garland Publishing, Inc, New York, NY, USA.

4. Kloas W and Lutz I. (2006). Amphibians as Model to Study Endocrine Disrupters. Journal of Chromatography A 1130: 16-27.

5. Chang C, Witschi E. (1956). Genic Control and Hormonal Reversal of Sex Differentiation in *Xenopus*. Journal of the Royal Society of Medicine 93: 140-144.

6. Gallien L. (1953) Total Inversion of Sex in *Xenopus laevis* Daud, Following Treatment with Estradiol Benzoate Administered During Larval Stage. Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences 237: 1565.

7. Villalpando I and Merchant-Larios H. (1990). Determination of the Sensitive Stages for Gonadal Sex-Reversal in *Xenopus Laevis* Tadpoles. International Journal of Developmental Biology 34. 281-285.

8. Miyata S, Koike S and Kubo T. (1999). Hormonal Reversal and the Genetic Control of Sex Differentiation in *Xenopus*. Zoological Science 16: 335-340.

9. Mikamo K and Witschi E. (1963). Functional Sex-Reversal in Genetic Females of *Xenopus laevis*, Induced by Implanted Testes. Genetics 48: 1411.

10. Olmstead AW, Kosian PA, Korte JJ, Holcombe GW, Woodis K and Degitz SJ. (2009)a. Sex reversal of the Amphibian, *Xenopus tropicalis*, Following Larval Exposure to an Aromatase Inhibitor. Aquatic Toxicology 91: 143-150.

11. Yoshimoto S, Okada E, Umemoto H, Tamura K, Uno Y, Nishida-Umehara C, Matsuda Y, Takamatsu N, Shiba T and Ito M. (2008). A W-linked DM-Domain Gene, DM-W, Participates in Primary Ovary Development in *Xenopus Laevis*. Proceedings of the National Academy of Sciences of the United States of America 105: 2469-2474.

12. Olmstead AW, Korte JJ, Woodis KK, Bennett BA, Ostazeski S and Degitz SJ. (2009)b. Reproductive Maturation of the Tropical Clawed Frog: *Xenopus tropicalis*. General and Comparative Endocrinology 160: 117-123.

13. Tobias ML, Tomasson J and Kelley DB. (1998). Attaining and Maintaining Strong Vocal Synapses in Female *Xenopus laevis*. Journal of Neurobiology 37: 441-448.

14. Qin ZF, Qin XF, Yang L, Li HT, Zhao XR and Xu XB. (2007). Feminizing/Demasculinizing Effects of Polychlorinated Biphenyls on the Secondary Sexual Development of *Xenopus Laevis*. Aquatic Toxicology 84: 321-327.

15. Porter KL, Olmstead AW, Kumsher DM, Dennis WE, Sprando RL, Holcombe GW, Korte JJ,

Lindberg-Livingston A and Degitz SJ. (2011). Effects of 4-*Tert*-Octylphenol on *Xenopus Tropicalis* in a Long Term Exposure. Aquatic Toxicology 103: 159-169.

16. ASTM. (2002). Standard Guide for Conducting Acute Toxicity Tests on Test Materials with Fishes, Macroinvertebrates, and Amphibians. ASTM E729-96, Philadelphia, PA, USA.

17. OECD (1992). Ready Biodegradability Test (No. 301). OECD Guidelines for the Testing of Chemicals, Organisation for Economic Cooperation and Development, Paris.

18. OECD (2014). Ready Biodegradability - CO2 in sealed vessels (Headspace Test) (No. 310). OECD Guidelines for the Testing of Chemicals. Organisation for Economic Cooperation and Development, Paris.

19. Kahl MD, Russom CL, DeFoe DL and Hammermeister DE. (1999). Saturation Units for Use in Aquatic Bioassays. Chemosphere 39: 539-551.

20. Adolfsson-Erici M, Åkerman G, Jahnke A, Mayer P, McLachlan MS (2012). A flow-through passive dosing system for continuously supplying aqueous solutions of hydrophobic chemicals to bioconcentration and aquatic toxicity tests. Chemosphere, 86(6), pp.593-9.

21. OECD (2000). Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures Environment, Health and Safety Publications, Series on testing and assessment (No. 23.), Organisation for Economic Cooperation and Development, Paris.

22. Hutchinson TH, Shillabeer N, Winter MJ and Pickford DB. (2006). Acute and Chronic Effects of Carrier Solvents in Aquatic Organisms: A Critical Review Review. Aquatic Toxicology 76. 69–92.

23. ASTM (2004). Standard Guide for Conducting the Frog Embryo Teratogenesis Assay - *Xenopus* (FETAX) ASTM E1439 - 98, Philadelphia, PA, USA.

24. Read BT (2005). Guidance on the Housing and Care of the African Clawed Frog *Xenopus Laevis*. Royal Society for the Prevention of Cruelty to Animals (RSPCA), Horsham, Sussex, U.K., 84 pp.

25. OECD (2009). The Amphibian Metamorphosis Assay. OECD Guidelines for the Testing of Chemicals (No. 231.) OECD Guidelines for the Testing of Chemicals Organisation for Economic Cooperation and Development, Paris.

26. OECD. (2012b). Fish Short Term Reproduction Assay. Guidelines for the Testing of Chemicals (No. 229.). OECD Guidelines for the Testing of Chemicals. Organisation for Economic Cooperation and Development, Paris.

27. OECD (2011). Fish Sexual Development Test. OECD Guidelines for the Testing of Chemicals (No. 234). OECD Guidelines for the Testing of Chemicals. Organisation for Economic Cooperation and Development, Paris.

28. OECD (2013). Fish Embryo Acute Toxicity (FET) Test. Guidelines for the Testing of Chemicals (No. 236.). OECD Guidelines for the Testing of Chemicals. Organisation for Economic Cooperation and Development, Paris.

29. OECD (2007). Guidance Document on Amphibian Thyroid Histology.Environment, Health and Safety Publications, Series on Testing and Assessment (No. 82.) Organisation for Economic Cooperation and Development, Paris.

30. Grim KC, Wolfe M, Braunbeck T, Iguchi T, Ohta Y, Tooi O, Touart L, Wolf DC and Tietge J. (2009). Thyroid Histopathology Assessments for the Amphibian Metamorphosis Assay to Detect Thyroid-Active Substances, Toxicological Pathology 37: 415-424.

31. Luna LG and Coady K ( 2014). Identification of *X. laevis* Vitellogenin Peptide Biomarkers for Quantification by Liquid Chromatography Tandem Mass Spectrometry. Analytical and Bioanalytical Techniques 5: (3): 194.

32. OECD (2015). Guidance on histopathology techniques and evaluation. Environment, Health and Safety Publications, Series on Testing and Assessment (No.228), Organisation for Economic Cooperation and Development, Paris.

33. OECD (2006). Current Approaches in the Statistical Analysis of Ecotoxicity Data: A Guidance to Application. Environment, Health and Safety Publications, Series on testing and assessment (No.54 ), Organisation for Economic Cooperation and Development, Paris.

34. Hutchinson TH, Bögi C, Winter MJ, Owens JW, 2009. Benefits of the Maximum Tolerated Dose (MTD) and Maximum Tolerated concentration (MTC) Concept in Aquatic Toxicology. Aquatic Toxicology;91(3), pp197-202.

## ANNEX 1

### DEFINITIONS

**Apical endpoint**: Causing effect at population level.

**ELISA**: Enzyme-Linked Immunosorbent Assay

**ECx: (**Effect concentration for x% effect) is the concentration that causes an x% of an effect on test organisms within a given exposure period when compared with a control. For example, an EC50 is a concentration estimated to cause an effect on a test end point in 50% of an exposed population over a defined exposure period.

**dpf**: days post fertilization

**Flow-through test**: is a test with continued flow of test solutions through the test system during the duration of exposure.

**HPG axis**: hypothalamic-pituitary-gonadal axis

**IUPAC**: International Union of Pure and Applied Chemistry.

**Lowest observed effect concentration (LOEC)** is the lowest tested concentration of a test chemical at which the chemical is observed to have a statistically significant effect (at $p < 0.05$) when compared with the control. However, all test concentrations above the LOEC should have a harmful effect equal to or greater than those observed at the LOEC. When these two conditions cannot be satisfied, a full explanation should be given for how the LOEC (and hence the NOEC) has been selected. ANNEX 7 provide guidance.

**Median Lethal Concentration (LC50)**: is the concentration of a test chemical that is estimated to be lethal to 50% of the test organisms within the test duration.

**No observed effect concentration (NOEC)** is the test concentration immediately below the LOEC, which when compared with the control, has no statistically significant effect ($p < 0.05$), within a stated exposure period.

**SMILES**: Simplified Molecular Input Line Entry Specification.

**UVCB**: substances of unknown or variable composition, complex reaction products or biological materials.

**VTG**: vitellogenin is a phospholipoglycoprotein precursor to egg yolk protein that normally occurs in sexually active females of all oviparous species.

## ANNEX 2

### SOME CHEMICAL CHARACTERISTICS OF AN ACCEPTABLE DILUTION WATER

| Substance | Limit concentration |
|---|---|
| Particulate matter | 5 mg/L |
| Total organic carbon | 2 mg/L |
| Un-ionised ammonia | 1 µg/L |
| Residual chlorine | 10 µg/L |
| Total organophosphorous pesticides | 50 ng/L |
| Total organochlorine pesticides plus polychlorinated biphenyls | 50 ng/L |
| Total organic chlorine | 25 ng/L |
| Aluminium | 1 µg/L |
| Arsenic | 1 µg/L |
| Chromium | 1 µg/L |
| Cobalt | 1 µg/L |
| Copper | 1 µg/L |
| Iron | 1 µg/L |
| Lead | 1 µg/L |
| Nickel | 1 µg/L |
| Zinc | 1 µg/L |
| Cadmium | 100 ng/L |
| Mercury | 100 ng/L |
| Silver | 100 ng/L |

## ANNEX 3

### TEST CONDITIONS FOR THE LAGDA

| | | |
|---|---|---|
| 1. | Test species | *Xenopus laevis* |
| 2. | Test type | Continuous flow-through, |
| 3. | Water temperature | The nominal temperature is 21 °C. The mean temperature over the duration of the test is 21 ± 1 °C (the inter-replicate and the inter-treatment differentials should not exceed 1.0 °C) |
| 4. | Illumination quality | Fluorescent bulbs (wide spectrum)<br>600-2000 lux (lumens/$m^2$) at the water surface |
| 5. | Photoperiod | 12 h light 12 h dark |
| 6. | Test solution volume and test vessel (tank) | 4-10 L (minimum 10–15 cm water depth)<br>Glass or stainless steel tank |
| 7. | Volume exchanges of test solutions | Constant, in consideration of both the maintenance of biological conditions and chemical exposure (*e.g.*, 5 tank volume renewal per day) |
| 8. | Age of test organisms at initiation | Nieuwkoop and Faber (NF) stage 8-10 |
| 9. | Number of organisms per replicate | 20 animals (embryos)/tank (replicate) at exposure initiation and 10 animals (juveniles)/tank (replicate) after NF stage 66 to exposure termination |
| 10. | Number of treatments | Minimum 4 test chemical treatments plus appropriate control(s) |
| 11. | Number of replicates per treatment | 4 replicates per treatment for test chemical and 8 replicates for control(s) |
| 12. | Number of organisms per test concentration | Minimum 80 animals per treatment for test chemical and minimum 160 replicates for control(s) |
| 13. | Dilution water | Any water that permits normal growth and development of *X. laevis* (*e.g.*, spring water or charcoal-filtered tap water) |
| 14. | Aeration | None required, but aeration of the tanks may be necessary if dissolved oxygen levels drop below recommended limits and increases in flow of test solution is maximized |
| 15. | Dissolved oxygen of test solution | Dissolved oxygen: ≥ 40 % of air saturation value or ≥ 3.5 mg/L |
| 16. | pH of test solution | 6.5-8.5 (the inter-replicate and the inter-treatment differentials should not exceed 0.5) |
| 17. | Hardness and alkalinity of test solution | 10-250 mg $CaCO_3$/L |

| | | |
|---|---|---|
| 18 | Feeding regime | (See ANNEX 4) |
| 19 | Exposure period | From NF stage 8-10 to ten weeks after the median time to NF stage 62 in water and/or solvent control group (maximum 17 weeks) |
| 20 | Biological endpoints | Mortality (and abnormal appearances), time to NF stage 62 (larval sample), thyroid histology assessment (larval sample), growth (weight and length), liver-somatic index (juvenile sample), genetic/phenotypic sex ratios (juvenile sample), histopathology for gonads, reproductive ducts, kidney and liver (juvenile sample) and plasma vitellogenin (juvenile sample, optional) |
| 21 | Test validity criteria | Dissolved oxygen should be > 40% air saturation value; mean water temperature should be 21 ± 1 °C and the inter-replicate and -treatment differentials should be < 1.0 °C, pH of test solution should be ranged between 6.5 and 8.5, the mortality in control should be ≤ 20% in each replicate , and the mean time to NF stage 62 in control should be ≤ 45 days; the mean weight of test organisms at NF stage 62 and at the termination of the assay in controls and solvent controls (if used) should reach 1.0 ± 0.2 and 11.5 ± 3 g, respectively, evidence should be available to demonstrate that the concentrations of the test chemical in solution have been satisfactorily maintained within ± 20% of the mean measured values |

## ANNEX 4

### FEEDING REGIME

It should be noted that although this feeding regime is recommended, alternatives are permissible providing the test organisms grow and develop at an appropriate rate.

***Larval feeding***

Preparation for larval diet

A. 1:1 (v/v) Trout Starter: algae/TetraFin® or equivalent ;
   1. Trout Starter: blend 50 g of Trout Starter (fine granules or powder) and 300 mL of suitable filtered water on a high blender setting for 20 seconds
   2. Algae/TetraFin® (or equivalent) mixture: blend 12 g spirulina algae disks and 500 ml filtered water on a high blender setting for 40 seconds, blend 12 g Tetrafin® with 500 ml filtered water and then combine these to make up 1 L of 12 g/L spirulina algae and 12 g/L Tetrafin®
   3. Combine equal volumes of the blended Trout Starter and the algae/TetraFin® mixture

B. Brine shrimp:

   15 ml brine shrimp eggs are hatched in 1 L of salt water (prepared by adding 20 mL of NaCl to 1 L deionized water). After aerating 24 hours at room temperature under constant light, the brine shrimp are harvested. Briefly, the brine shrimp are allowed to settle for 30 min by stopping aeration. Cysts that float to the top of the canister are poured off and discarded, and the shrimp are poured through the appropriate filters and brought up to 30 ml with filtered water.

*Feeding Protocol*

**Table 1** provides a reference regarding the type and amount of feed used during the larval stages of the exposure. The animals should be fed three times per day Monday through Friday and once per day on the weekends.

**Table 1.** Feeding regime for *X. laevis* larvae in flow-through conditions.

| Time* (Post Fertilization) | Trout Starter: algae/TetraFin® | | Brine Shrimp | |
|---|---|---|---|---|
| | Weekday (3 times per day) | Weekend (once per day) | Weekday (twice per day) | Weekend (once per day) |
| Days 4-14 (in Weeks 0-1) | 0.33 ml | 1.2 ml | 0.5 ml (from Day 8 to 15) | 0.5 ml (from Day 8 to 15) |
| Week 2 | 0.67 ml | 2.4 ml | 1 ml (from Day 16) | 1 ml (from Day 16) |
| Week 3 | 1.3 ml | 4.0 ml | 1 ml | 1 ml |
| Week 4 | 1.5 ml | 4.0 ml | 1 ml | 1 ml |
| Week 5 | 1.6 ml | 4.4 ml | 1 ml | 1 ml |
| Week 6 | 1.6 ml | 4.6 ml | 1 ml | 1 ml |
| Week 7 | 1.7 ml | 4.6 ml | 1 ml | 1 ml |
| Weeks 8-10 | 1.7 ml | 4.6 ml | 1 ml | 1 ml |

* Day 0 is defined as the day hCG injection is done.

***Larval to juvenile diet transition***

As larvae complete metamorphosis, they transition to a juvenile diet formulation explained below. While this transition is taking place, the larval diet should be reduced as the juvenile feed increases. This can be accomplished by proportionally decreasing the larval feed while proportionally increasing the juvenile feed as each group of five tadpoles surpass NF stage 62 and approach completion of metamorphosis at NF stage 66.

***Juvenile feeding***

*Juvenile diet*

Once metamorphosis is complete (stage 66), the feeding regime changes to 3/32 inch premium sinking frog food alone (Xenopus Express, FL, USA), or equivalent.

*Preparation of crushed pellet for larval to juvenile transition*

Sinking frog food pellets are briefly run in a coffee grinder, blender or mortar and pestle in order to reduce the size of the pellets by approximately 1/3. Processing too long results in powder and is discouraged.

*Feeding protocol*

**Table 2** provides a reference regarding the type and amount of feed used during juvenile and adult life stages. The animals should be fed once per day. It should be noted that as animals metamorphose, they continue receiving a portion of the brine shrimp until > 95% animals complete metamorphosis.

The animals should not be fed on the day of test termination so feed does not confound weight measurements.

**Table 2.** Feeding regime for *X. laevis* juveniles in flow-through conditions. It should be noted that unmetamorphosed animals, including those whose metamorphosis has been delayed by the chemical treatment, cannot eat uncrushed pellets.

| Time (Weeks post-median metamorphosis date) | Crushed pellet volume (mg per froglet) | Whole pellet volume (mg per froglet) |
|---|---|---|
| As animals complete metamorphosis | 25 | 0 |
| Weeks 0-1 | 25 | 28 |
| Weeks 2-3 | 0 | 110 |
| Weeks 4-5 | 0 | 165 |
| Weeks 6-9 | 0 | 220 |

* The first day of Week 0 is the median metamorphosis date in control animals.

ANNEX 5

GENETIC SEX DETERMINATION (GENETIC SEXING)

The method of genetic sexing for *Xenopus laevis* is based on Yoshimoto *et al.*, 2008. Procedures in detail on the genotyping can be obtained from this publication, if needed. Alternative methods (e.g. high-throughput qPCR) may be used if considered suitable.

***X. laevis primers***

DM-W marker
Forward: 5'-CCACACCCAGCTCATGTAAAG-3'
Reverse: 5'-GGGCAGAGTCACATATACTG-3'

Positive Control
Forward: 5'-AACAGGAGCCCAATTCTGAG-3'
Reverse: 5'-AACTGCTTGACCTCTAATGC-3'

***DNA purification***

Purify DNA from muscle or skin tissue using e.g., Qiagen DNeasy Blood and Tissue Kit (cat # 69506) or similar product according to kit instructions. DNA can be eluted from the spin columns using less buffer to yield more concentrated samples if deemed necessary for PCR. Note that DNA is quite stable, so care should be taken to avoid cross-contamination that could lead to mischaracterization of males as females, or vice versa.

***PCR***

A sample protocol using JumpStart™ *Taq* from Sigma is outlined in **Table 1**.

**Table 1** Sample *protocol* using JumpStart™ *Taq* from Sigma

| Master Mix | 1x (µL) | [Final] |
|---|---|---|
| NFW | 11 | - |
| 10X Buffer | 2.0 | - |
| $MgCl_2$ (25mM) | 2.0 | 2.5 mM |
| dNTP's (10mM each) | 0.4 | 200 µM |
| Marker for primer (8 µM) | 0.8 | 0.3 µM |
| Marker rev primer (8 µM) | 0.8 | 0.3 µM |
| Control for primer (8 µM) | 0.8 | 0.3 µM |
| Control rev primer (8 µM) | 0.8 | 0.3 µM |
| JumpStart™ Taq | 0.4 | 0.05 units/µl |
| DNA template | 1.0 | ~200 pg/µl |

Note: When preparing Master Mixes, prepare extra to account for any loss that may occur while pipetting (example: 25x should be used for only 24 reactions).

Reaction:

Master Mix 19.0 µL
Template 1.0 µL
Total 20.0 µL

Thermocycler Profile:

Step 1. 94 °C 1 min
Step 2. 94 °C 30 sec
Step 3. 60 °C 30 sec
Step 4. 72 °C 1 min
Step 5. Go to step 2. 35 cycles
Step 6. 72 °C 1 min
Step 7. 4 °C hold

PCR products can be run immediately in a gel or stored at 4 °C.

*Agarose Gel Electrophoresis (3%)*

50X TAE

Tris 24.2 g
Glacial acetic acid 5.71 mL
$Na_2$ (EDTA)·$2H_2O$ 3.72 g
Add water to 100 mL

1X TAE

$H_2O$ 392 mL
50X TAE 8 mL

3:1 Agarose

3 parts NuSieve™ GTG™ agarose
1 part Fisher agarose low electroendosmosis (EEO)

*Method*

1. Prepare a 3% gel by adding 1.2 g agarose mix to 43 mL 1X TAE. Swirl to disassociate large clumps.
2. Microwave agarose mixture until completely dissolved (avoid boiling over). Let cool slightly.
3. Add 1.0 µL ethidium bromide (10 mg/mL). Swirl flask. Note that ethidium bromide is mutagenic, so alternative chemicals could be used for this step to minimize health risks to workers.
4. Pour gel into mold with comb. Cool completely.
5. Add gel to apparatus. Cover gel with 1X TAE.
6. Add 1 µL of 6x loading dye to each 10 µl PCR product.

7. Pipette samples into wells.
8. Run at 160 constant volts for ~20 minutes.

An agarose gel image showing the band patterns indicative of male and female individuals is shown in **Figure 1**.

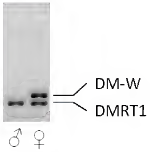


**Figure 1.**Agarose gel image showing the band pattern indicative of a male (♂) individual (single band ~203 bp: DMRT1) and of a female (♀) individual (two bands at ~259 bp: DM-W and 203 bp:DMRT1).

## Literature

Yoshimoto S, Okada E, Umemoto H, Tamura K, Uno Y, Nishida-Umehara C, Matsuda Y, Takamatsu N, Shiba T, Ito M. 2008. A W-linked DM-domain gene, DM-W, participates in primary ovary development in *Xenopus laevis*. Proceedings of the National Academy of Sciences of the United States of America 105: 2469-2474.

## ANNEX 6

### MEASUREMENT OF VITELLOGENIN

The measurement of vitellogenin (VTG) is made using an enzyme-linked immunosorbent assay (ELISA) method which was originally developed for fathead minnow VTG (Parks *et al.*, 1999). Currently there are no commercially available antibodies for *X. laevis*. However, given the wealth of information for this protein and the availability of cost-effective commercial antibody production services, it is reasonable that laboratories can easily develop an ELISA to make this measure (Olmstead *et al.*, 2009). Also Olmstead *et al.* (2009) provide a description of the assay as modified for VTG in *X. tropicalis*, as shown below. The method uses an antibody made against *X. tropicalis* VTG, but it is known also to work for *X. laevis* VTG. It should be noted that non-competitive ELISAs can also be used, and that these may have lower detection limits than the method described below.

#### *Materials and Reagents*

- Preadsorbed 1st Antibody (Ab) serum

  Mix 1 part anti-*X. tropicalis* VTG 1st Ab serum with 2 parts control male plasma and leave at RT for ~ 75 minutes, put on ice for 30 min, centrifuge > 20K x G for 1 hour at 4 °C, remove supernatant, aliquot, store at -20 °C.

- 2nd Antibody

  Goat Anti-Rabbit IgG-HRP conjugate (e.g., Bio-Rad 172-1019)

- VTG Standard

  purified *X. laevis* VTG at 3.3 mg/ml.

- TMB (3,3',5,5' Tetramethyl-benzidine) (e.g., KPL 50-76-00; or Sigma T0440)
- Normal Goat Serum (NGS) (e.g., Chemicon S26-100ml)
- 96 well EIA polystyrene microtiter plates (e.g., ICN. 76-381-04, Costar:53590, Fisher:07-200-35)
- 37 °C hybridization oven (or fast equilibrating air incubator) for plates, water bath for tubes
- Other common laboratory equipment, chemicals, and supplies.

#### *Recipes*

- Coating Buffer (50 mM Carbonate Buffer, pH 9.6):

  $NaHCO_3$ 1.26 g
  $Na_2CO_3$ 0.68 g
  water 428 ml

- 10X PBS (0.1 M phosphate, 1.5 M NaCl):

  $NaH_2PO_4{\cdot}H_2O$ 0.83 g
  $Na_2HPO_4{\cdot}7\,H_2O$ 20.1 g
  NaCl 71 g
  water 810 ml

- Wash Buffer (PBST):

  10X PBS 100 ml
  water 900 ml

Adjust pH to 7.3 with 1 M HCL, then add 0.5 ml Tween-20

- Assay Buffer:

  Normal Goat Serum (NGS) 3.75 ml
  Wash Buffer 146.25 ml

***Sample collection***

Blood is collected with an heparinized microhematocrit tube and placed on ice. After centrifugation for 3 minutes, the tube is scored, broken open, and the plasma expelled into 0.6 ml microcentrifuge tubes which contain 0.13 units of lyophilized aprotinin. (These tubes are prepared in advance by adding the appropriate amount of aprotinin, freezing, and lyophilizing in a speed-vac at low heat until dry.) Store plasma at -80 °C until analyzed.

***Procedure for one plate***

*Coating the plate*

Mix 20 µl of purified VTG with 22 ml of carbonate buffer (final 3 µg/ml). Add 200 µl to each well of a 96-well plate using. Cover the plate with adhesive sealing film and allow to incubate overnight at 37 °C for 2 hours (or 4 °C overnight).

*Blocking the plate*

Blocking solution is prepared by adding 2 ml of Normal Goat Serum (NGS) to 38 ml of carbonate buffer. Remove coating solution and shake dry. Add 350 µl of the blocking solution to each well. Cover with adhesive sealing film and incubate at 37 °C for 2 hours (or at 4 °C overnight).

*Preparation of standards*

5.8 µl of purified VTG standard is mixed with 1.5 ml of assay buffer in a 12 x 75 mm borosilicate disposable glass test tube. This yields 12,760 ng/ml. Then a serial dilution is performed by adding 750 µl of the previous dilution to 750 µl of assay buffer to yield final concentrations of 12,760, 6380, 3190, 1595, 798, 399, 199, 100, and 50 ng/ml.

*Preparation of Samples*

Start with a 1:300 (e.g., combine 1 µl plasma with 299 µl of assay buffer) or 1:30 dilution of plasma into assay buffer. If a large amount of VTG is expected, additional or greater dilutions may be needed. Try to keep $B/B_o$ within the range of standards. For samples without appreciable VTG, e.g., control males and females (which are all immature), use the 1:30 dilution. Samples diluted less than this may show unwanted matrix effects.

Additionally, it is recommended to run a positive control sample on each plate. This comes from a pool of plasma containing highly induced levels of VTG. The pool is initially diluted in NGS, divided in aliquots and stored at -80 C. For each plate, an aliquot is thawed, diluted further in assay buffer and run similar to a test sample.

*Incubation with 1st antibody*

The 1st Ab is prepared by making a 1:2,000 dilution of preadsorbed 1st Ab serum in assay buffer (e.g., 8 µl to 16 ml of assay buffer). Combine 300 µl of 1st Ab solution with 300 µl of sample/standard in a glass

tube. The $B_o$ tube is prepared similarly with 300 µl of assay buffer and 300 µl of antibody. Also, a NSB tube should be prepared using 600 µl of assay buffer only, i.e., no Ab. Cover the tubes with Parafilm and vortex gently to mix. Incubate in a 37 °C water bath for 1 hour.

*Washing the plate*

Just before the 1st Ab incubation is complete, wash the plate. This is done by shaking out the contents and patting dry on absorbent paper. Then fill wells with 350 µl of wash solution, dump out, and pat dry. A multi-channel repeater pipette or plate washer is useful here. The wash step is repeated two more times for a total of three washes.

*Loading the plate*

After the plate has been washed, remove the tubes from the water bath and vortex lightly. Add 200 µl from each sample, standard, $B_o$, and NSB tube to duplicate wells of the plate. Cover plate with adhesive sealing film and allow to incubate for 1 hour at 37 °C.

*Incubation with the 2nd antibody*

At the end of the incubation from the previous step, the plate should be washed three times again, like above. The diluted 2nd Ab is prepared by mixing 2.5 µl of 2nd Ab with 50 ml of assay buffer. Add 200 µl of diluted 2nd Ab to each well, seal like above, and incubate for 1 hour at 37 °C.

*Addition of substrate*

After the incubation with the 2nd Ab is complete, wash the plate three times as described earlier. Then add 100 µl of TMB substrate to each well. Allow the reaction to proceed for 10 minutes, preferably out of bright light. Stop the reaction by adding 100 µl of 1 M phosphoric acid. This will change the color from blue to an intense yellow. Measure the absorbance at 450 nm using a plate reader.

***Calculate B/Bo***

Subtract the average NSB value from all measurements. The $B/B_o$ for each sample and standard is calculated by dividing the absorbance value (B) by the average absorbance of the $B_o$ sample.

***Obtain the standard curve and determine unknown amounts***

Generate a standard curve with the aid of some computer graphing software (e.g., Slidewrite or Sigma Plot®) that will extrapolate quantity from $B/B_o$ of sample based on $B/B_o$ of standards. Typically, the amount is plotted on a log scale and the curve has a sigmoid shape. However, it may appear linear when using a narrow range of standards. Correct sample amounts for dilution factor and report as mg VTG/ml of plasma.

***Determination of minimum detection limits (MDL)***

Often, particularly in normal males, it will not be clear how to report results from low values. In these cases, the 95% "Confidence limits" should be used to determine if the value should be reported as zero or as some other number. If the sample result is within the confidence interval of the zero standard ($B_o$), the result should be reported as zero. The minimum detection level will be the lowest standard which is consistently different from the zero standard; that is, the two confidence intervals don't overlap. For any sample result which is within the confidence limit of the minimum detection level, or above, the calculated value will be reported. If a sample falls between the zero standard and the minimum detection level

intervals, one half of the minimum detection level should be reported for the value of that sample.

**Literature**

Olmstead AW, Korte JJ, Woodis KK, Bennett BA, Ostazeski S, Degitz SJ. 2009. Reproductive maturation of the tropical clawed frog: *Xenopus tropicalis*. General and Comparative Endocrinology 160: 117-123.

Parks LG, Cheek AO, Denslow ND, Heppell SA, McLachlan JA, LeBlanc GA, Sullivan CV. 1999. Fathead minnow (*Pimephales promelas*) vitellogenin: purification, characterization and quantitative immunoassay for the detection of estrogenic compounds. Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology 123: 113-125.

## ANNEX 7

### STATISTICAL ANALYSIS

The LAGDA generates three forms of data to be statistically analyzed: (1) Quantitative continuous data, (2) Time-to-event data for developmental rates (Time to NF stage 62) and (3) Ordinal data in the form of severity scores or developmental stages from histopathology evaluations. The recommended statistical analysis decision tree for the LAGDA is shown in **Figure 1**. Also, some annotations which might be needed to conduct statistical analysis for the measurements from the LAGDA are indicated below. For the analysis decision tree, the results of measurements for mortality, growth (weight and length) and liver-somatic-index (LSI) should be analyzed according to the "Other endpoints" branch.

***Continuous data***

Data for continuous endpoints should first be checked for monotonicity by rank transforming the data, fitting to an ANOVA model and comparing linear and quadratic contrasts. If the data are monotonic, a step-down Jonckheere-Terpstra trend test should be performed on replicate medians and no subsequent analyses should be applied. An alternative for data that are normally distributed with homogeneous variances is the step-down Williams' test. If the data are non-monotonic (quadratic contrast is significant and linear is not significant), they should be analyzed using a mixed effects ANOVA model. The data should then be assessed for normality (preferably using the Shapiro-Wilk or Anderson-Darling test) and variance homogeneity (preferably using Levene's test). Both tests are performed on the residuals from the mixed effects ANOVA model. Expert judgment can be used in lieu of these formal tests for normality and variance homogeneity, though formal tests are preferred. If the data are normally distributed with homogeneous variance, then the assumptions of a mixed effect ANOVA are met and a significant treatment effect is determined from Dunnett's test. Where non-normality or variance heterogeneity is found, then the assumptions of Dunnett's test are violated and a normalizing, variance stabilizing transform is sought. If no such transform is found, then a significant treatment effect is determined with a Dunn's test. Whenever possible, a one-tailed test should be performed as opposed to a two-tailed test, but it requires expert judgment to determine which is appropriate for a given endpoint.

*Mortality*

Mortality data should be analyzed for the time period encompassing the full test and should be expressed as proportion that died in any particular tank. Tadpoles that do not complete metamorphosis in the given time frame, those tadpoles that are in the larval sub-sample cohort, those juvenile frogs that are culled, and any animal that dies due to experimenter error should be treated as censored data and not included in the denominator of the percent calculation. Prior to any statistical analyses, mortality proportions should be arcsin-square root transformed. An alternative is to use the step-down Cochran-Armitage test, possibly with a Rao-Scott adjustment in the presence of overdispersion.

*Weight and length (growth data)*

Males and females are not sexually-dimorphic during metamorphosis so larval sub-sampling growth data should be analyzed independent of gender. However, juvenile growth data should be analyzed separately based on genetic sex. A log-transformation may be needed for these endpoints since log-normality of size data is not uncommon.

*Liver-somatic-index (LSI)*

Liver weights should be normalized as proportions of whole body weights (i.e., LSI) and analyzed separately based on genetic sex.

***Time to NF stage 62***

Time to metamorphosis data should be treated as time-to-event data, with any mortalities or individuals not reaching NF stage 62 in 70 days treated as right-censored data (i.e. the true value is greater than 70 days but the study ends before the animals had reached NF stage 62 in 70 days). Median time to NF stage 62 completion of metamorphosis in dilution water controls should be used to determine the test termination date. Median time to completion of metamorphosis could be determined by Kaplan-Meier product-limit estimators. This endpoint should be analyzed using a mixed-effects Cox proportional hazard model that takes account of the replicate structure of the study.

***Histopathology data (severity scores and developmental stages)***

Histopathology data are in the form of severity scores or developmental stages. A test termed RSCABS (Rao-Scott Cochran-Armitage by Slices) uses a step-down Rao-Scott adjusted Cochran-Armitage trend test on each level of severity in a histopathology response (Green *et al.*, 2014). The Rao-Scott adjustment incorporates the replicate vessel experimental design into the test. The "by Slices" procedure incorporates the biological expectation that severity of effect tends to increase with increasing doses or concentrations, while retaining the individual subject scores and revealing the severity of any effect found. The RSCABS procedure not only determines which treatments are statistically different from controls (i.e., have more severe pathology than controls), but it also determines at which severity score the difference occurs thereby providing much needed context to the analysis. In the case of developmental staging of gonads and reproductive ducts, an additional manipulation should be applied to the data since an assumption of RSCABS is that severity of effect increases with dose. The effect observed could be a delay or acceleration of development. Therefore, developmental staging data should be analyzed as reported to detect acceleration in development and then manually inverted prior to a second analysis to detect a delay in development.

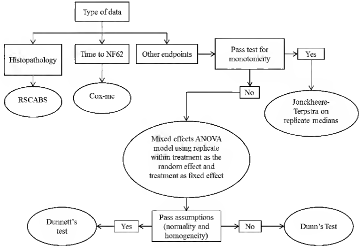


**Figure 1.**Statistical analysis decision tree for LAGDA data.

**Literature**

Green JW, Springer TA, Saulnier AN, Swintek J. 2014. Statistical analysis of histopathology endpoints. Environmental Toxicology and Chemistry 33, 1108-1116.

## ANNEX 8

### CONSIDERATIONS FOR TRACKING AND MINIMIZING THE OCCURRENCE OF SCOLIOSIS

Idiopathic scoliosis, usually manifesting as "bent tail" in *Xenopus laevis* tadpoles, may complicate morphological and behavioral observations in test populations. Efforts should be made to minimize or eliminate the incidence of scoliosis, both in stock and under test conditions. In the definitive test, it is recommended that the prevalence of moderate and severe scoliosis be less than 10%, to improve confidence that the test can detect treatment-related developmental effects in otherwise healthy amphibian larvae.

Daily observations during the definitive test should record both the incidence (individual count) and severity of scoliosis, when present. The nature of the abnormality should be described with respect to location (*e.g.*, anterior or posterior to the vent) and direction of curvature (*e.g.*, lateral or dorsal-to-ventral). Severity may be graded as follows:

(NR) Not remarkable: no curvature present

(1) Minimal: slight, lateral curvature posterior to the vent; apparent only at rest

(2) Moderate: lateral curvature posterior to the vent; visible at all times but does not inhibit movement

(3) Severe: lateral curvature anterior to the vent; OR any curvature that inhibits movement; OR any dorsal-to-ventral curvature

A US EPA FIFRA Scientific Advisory Panel (FIFRA SAP 2013) reviewed summary data for scoliosis in fifteen Amphibian Metamorphosis Assays with *X. laevis* (NF stage 51 through 60+) and provided general recommendations for reducing the prevalence of this abnormality in test populations. The recommendations are relevant to the LAGDA even though this test encompasses a longer developmental timeline.

#### ***Historical Spawning Performance***

Generally, high quality, healthy adults should be used as breeding pairs; eliminating breeding pairs that produce offspring with scoliosis may minimize its occurrence over time. Specifically, minimizing the use of wild-caught breeding stock may be beneficial. The LAGDA exposure period begins with NF stage 8-to-10 embryos, and it is not feasible to determine at the test outset whether given individuals will exhibit scoliosis. Thus, in addition to tracking the incidence of scoliosis in animals that are placed on test, historical clutch performance (including the prevalence of scoliosis in any larvae allowed to develop) should be documented. It may be useful to further monitor the portion of each clutch not used in a given study and to report these observations (FIFRA SAP 2013).

#### ***Water Quality***

It is important to ensure adequate water quality, both in laboratory stock and during the test. In addition to water quality criteria routinely evaluated for aquatic toxicity tests, it may be useful to monitor for and to correct any nutrient deficiencies (*e.g.*, deficiency of vitamin C, calcium, phosphorus) or excess levels of selenium and copper, which are reported to cause scoliosis to varying degrees in laboratory-reared *Rana* sp. and *Xenopus* sp. (Marshall *et al.* 1980; Leibovitz *et al.* 1992; Martinez *et al.* 1992; as reported in FIFRA SAP 2013). The use of an appropriate dietary regimen (see ANNEX 4), and regular tank cleaning, will generally improve water quality and health of the test specimens.

***Diet***

Specific recommendations for a dietary regimen, found to be successful in the LAGDA, are detailed in ANNEX 4. It is recommended that feed sources be screened for biological toxins, herbicides, and other pesticides which are known to cause scoliosis in *X. laevis* or other aquatic animals (Schlenk and Jenkins 2013). For example, exposure to certain cholinesterase inhibitors has been associated with scoliosis in fish (Schultz *et al.* 1985) and frogs (Bacchetta *et al.* 2008).

***References***

Bacchetta, R., P. Mantecca, M. Andrioletti, C. Vismara, and G. Vailati. 2008. Axial-skeletal defects caused by carbaryl in *Xenopus laevis* embryos. Science of the Total Environment 392: 110 – 118.

Schultz, T.W., J.N. Dumont, and R.G. Epler. 1985. The embryotoxic and osteolathyrogenic effects of semicarbazide. Toxicology 36: 185-198.

Leibovitz, H.E., D.D. Culley, and J.P. Geaghan. 1982. Effects of vitamin C and sodium benzoate on survival, growth and skeletal deformities of intensively culture bullfrog larvae (*Rana catesbeiana*) reared at two pH levels. Journal of the World Aquaculture Society 13: 322-328.

Marshall, G.A., R.L. Amborski, and D.D. Culley. 1980. Calcium and pH requirements in the culture of bullfrog (*Rana catesbeiana*) larvae. Journal of the World Aquaculture Society 11: 445-453.

Martinez, I., R. Alvarez, I. Herraez, and P. Herraez. 1992. Skeletal malformations in hatchery reared *Rana perezi* tadpoles. Anatomical Records 233(2): 314-320.

Schlenk, D., and Jenkins, F. 2013. Endocrine Disruptor Screening Prog (EDSP) Tier 1 Screening Assays and Battery Performance. US EPA FIFRA SAP Minutes No. 2013-03. May 21-23, 2013. Washington, DC.